ABSTRACT

This thesis focuses on four main research themes namely shot boundary detection, fast frame alignment, activity-driven video summarisation, and highlights based video annotation and retrieval. A number of novel algorithms have been proposed to address these issues, which can be highlighted as follows.

Firstly, accurate and robust shot boundary detection is achieved through modelling of cuts into sub-categories and appearance based modelling of several gradual transitions, along with some novel features extracted from compressed video. Secondly, fast and robust frame alignment is achieved via the proposed subspace phase correlation (SPC) and an improved sub-pixel strategy. The SPC is proved to be insensitive to zero-mean-noise, and its gradient-based extension is even robust to non-zero-mean noise and can be used to deal with non-overlapped regions for robust image registration. Thirdly, hierarchical modelling of rush videos using formal language techniques is proposed, which can guide the modelling and removal of several kinds of junk frames as well as adaptive clustering of retakes. With an extracted activity level measurement, shot and sub-shot are detected for content-adaptive video summarisation. Fourthly, highlights based video annotation and retrieval is achieved, in which statistical modelling of skin pixel colours, knowledge-based shot detection, and improved determination of camera motion patterns are employed.

Within these proposed techniques, one important principle is to integrate various kinds of feature evidence and to incorporate prior knowledge in modelling the given problems. High-level hierarchical representation is extracted from the original linear structure for effective management and content-based retrieval of video data. As most of the work is implemented in the compressed domain, one additional benefit is the achieved high efficiency, which will be useful for many online applications.

DEDICATION

吾生也有涯，而知也无涯

Our life is limited, but the knowledge

to explore is boundless.

—庄子(from Chuang-tzu, 369BC-286BC)

This thesis is dedicated to my wife Yujing,

my son Tianqi and my Parents.

Without their constant love, encouragement and support,

this work would have no chance been possible.

ACKNOWLEDGEMENTS

First of all, I would like thank Prof. J. Jiang and Dr. S. S. Ipson for their valuable guidance and ever support throughout my PhD study. The relatively loose research environment has brought me great freedom in exploring several different topics and facing the relevant challenges. Numerous discussions on all types of research problems with them have led me to this success.

I would also like to thank all my colleagues and friends in Bradford, especially Miss Juan Chen, Ms Chunmei Qing, and Dr. Ying Weng for the support and many useful discussions with them. Special thanks are owing to Miss Rona Wilson for her kind support.

Though having been dedicated before, I have to again express my grateful appreciation to my family, especially to my wife Yujing as well as my son, my parents and others. It is their strong support and encouragement that has helped me to complete this study.

I am very grateful for Prof. H. Hu at Essex University and Dr. D. Rigas to be my external and internal examiners. Their comments and suggestions have improved the quality of this thesis and made it more fluent and readable.

Finally, I wish to acknowledge the financial support from the EU IST FP6 Project "LIVE" (Contract No. IST-4-027312) and also thank the TRECVID community for the wonderful platform they have provided for me to develop and evaluate the proposed techniques.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES